TECHNICAL REPORT

SOIL MOISTURE IN KAZAKHSTAN:

IN SITU PROBES AND SATELLITE DATA

JULY 2014

This publication is made possible by the support of the American people through the United States Agency for International Development (USAID). It was prepared by Engility/International Resources Group (IRG) and WeatherPredict Consulting Inc.

This report has been prepared for the United States Agency for International Development (USAID), under the Climate Change Resilient Development Task Order No. AID-OAA-TO-11-00040, under The Integrated Water and Coastal Resources Management Indefinite Quantity Contract (WATER IQC II) Contract No. AID-EPP-I-00-04-00024.

Engility Corporation Contact:
Glen Anderson, Chief of Party, Glen.Anderson@EngilityCorp.com
Engility Corporation
1320 Braddock Place
Alexandria, VA 22314

Cover Photo: Daniel Byers, Skyship Films

SOIL MOISTURE IN KAZAKHSTAN:

IN SITU PROBES AND SATELLITE DATA

July 2014

Prepared for:

United States Agency for International Development

Global Climate Change Office, Climate Change Resilient Development project

Washington, DC

Prepared by:

Alan Basist
WeatherPredict Consulting Inc,
Raleigh, North Carolina

Svetlana Dolgih
Kazhydromet
Astana, Kazakhstan

Azamat Kutuzov
National Space Institute
Washington, D.C.

Neil Thomas
Resource Data
Asheville, North Carolina

Contact: Michael E. Cote, Engility Corporation, Michael.Cote@EngilityCorp.com

TABLE OF CONTENTS

ACRONYMS

BWI	Basist Wetness Index
PD	Probe Data
SAR	Synthetic Aperture Radar
SSMI	Special Sensor Microwave Imager
WI	Wetness Index

LIST OF FIGURES

1. ABSTRACT

Soil moisture is a critical factor affecting the production of wheat across northern Kazakhstan. Wheat yield in Kazakhstan is strongly related to food security and international stability in Central Asia. In an effort to understand and monitor how soil moisture affects yields, probe data was used to detect fluctuations, which relates to variability in the satellite derived wetness values. The soil moisture observations provided by probe measurements serve as calibration points to the satellite measurements. The regression equations derived from these relationships identify the covariance between the quantities. Many of the relationships between the probe data and wetness index were meaningful and can be utilized to effectively identify how upper level moisture fluctuates in Northern Kazakhstan during the period of wheat production. The slope and intercepts of the equations determine the ratio between the two measurements, and the intercept identifies when the soil is effectively dry, relative to the satellite observation. Generally, there were two unique relationships: one for the summer season (June, July, and August) and one for May, when there was nominal vegetation covering the surface.

Findings from this study were highly significant and can be applied in near real time in order to monitor the distribution of upper level soil moisture across the northern oblasts of Kazakhstan, where the production of wheat is critical to food security in Central Asia.

2. INTRODUCTION

The three northern oblasts of Kazakhstan are major sources of quality wheat to many countries of central Asia. Since food security is essential to national welfare and international cooperation across Central Asia (Weinthal, E. 2002), this study associates available soil moisture to the stability of the region. The extreme variability in weather and the potential impact of climate on water and yields has prompted concern that the region is vulnerable to climate change (Vörösmarty, et al. 2000). Specifically, a limiting factor of wheat yields is soil moisture in this sensitive region. Therefore, it is important to understand the amount of moisture in the soil throughout the region (Robinson et al. 2008), particularly during the active period of the growing season.

As a response to this need, the Kazakhstan government has established a considerable network of moisture probes in the oblasts of Northern Kazakhstan. These probes are located in Kostanay, Northern Kazakhstan, and Akmola (Figure 1). The Government of Kazakhstan uses information from regional probe data to help predict fluctuations in crop yields. Probe data (PD) is available in less than 50 sites through the region, and the data is not available in a digital format until months after the observations are made. Before data can be applied, it must be digitized and quality controlled, which can significantly delay its application and utility for assessing growing conditions in the region. These are some inherent strengths and weaknesses in the *in situ* PD used in the study.

Figure 1. A map of the study regions (8, 9, 10) where soil moisture measurements are taken.

Another approach is to use remotely sensed observations from space. Alsdorf et al. (2007) provides an overview of many of the techniques to do so and their applications. Griffiths and Wooding (1996) describe how Synthetic Active Radar (SAR) can be used to monitor temporal variability in soil moisture. This study utilizes a Wetness Index (WI), which is derived from the Special Sensor Microwave Imager (SSMI) (Basist et al 2001). These WI observations are available in near real time and are provided as weekly and monthly products. In this study, the PD is incorporated and correlated with SSMI surface wetness. The goal is to determine if there is a meaningful relationship between upper level soil moisture (as identified by the probe measurement) and the surface wetness index observed by the SSMI satellite instrument.

The study will use the values of the Wetness Index to predict probe data during the important months of the growing season: May, June, July, and August. **The hypothesis is that the relationship between the two measurements could be applied throughout the wheat growing area. This would allow the Kazhydromet to expand the distribution of observations from less than 50 to several hundred.** It would also provide near real time values while the probe network data may only be available months later.

3. METHODOLOGY

The primary goal of this study is to derive a linear relationship between probe data and satellite measurements. Therefore they must be provided as vectors over time. The PD is available on a ten day period and the measurement is usually taken on the 8th, 18th, and 28th of the month[1]. This allows up to three observations per month at each station. The period of observations began in 2003 and extended through 2012, although the data in the early half of the record tends to be sparse.

Probe data (PD) is provided from three oblasts: North Kazakhstan, Akmola, and Kostanay. The PD is used as the calibration source to simulate observations with the WI. We chose the coincide probe station that falls inside a 33 km by 33 km satellite pixel, and if the probe was near the boundary of the two satellite pixels, we included the average of multiple pixels against the probe measurements. While probe measurements are taken in ten day intervals, the WI combined up to 14 observations into a weekly observation. Therefore, it was determined to find an appropriate relationship between the two measurements and associate them over the period of record. A higher weighting of the satellite data was shown to have the best correspondence to the end of the ten day period.

Notwithstanding these adjustments, there remains a fundamental difference between the two observations. The significance of this discrepancy is minimal, as soils tend to have considerable lag response and memory of recent rain and/or snow melt events. This premise shows the two observations as complementary procedures that would have coinciding variability in their measurement of upper level soil moisture. It was speculated that there would be sufficient periods when the relationships would be apparent, a clear signal would be detected, and a statistical formula that identifies their relationship could be derived.

In order to extract these relationships, outliers and unreliable data were removed from the analysis. An investigation with personnel at the Kazhydromet and the National Space Institute of Kazakhstan demonstrated the benefit of removing outliers by using only reliable data to identify the relationship between the PD and WI. There was recognition that the two data sets would measure considerably different conditions based on the way their independent measurements are derived. Stable results were achieved by being liberal about which data to retain in the analysis.

For efficiency, a linear relationship between the corresponding data sets was assumed. Originally, data was parsed into independent relationships for each probe station. The stations that demonstrated meaningful relationships between the two data were retained in the subsequent analyses (see Figure 2).

[1] Measurements may be delayed a day or two if it is raining during scheduled observation times.

Probes used in the Regression Equations dureing the month of May

Oblast	Probe Stations	Oblast	Probe Stations	Oblast	Probe Stations
North Kazakhstan	Blagoveschenka	Akmola	Balkashino	Kostanay	Arshalinskiy
North Kazakhstan	Bulaevo	Akmola	Zhaltyr	Kostanay	Diyevskaya
North Kazakhstan	Chkalovo	Akmola	Shortandy	Kostanay	Karasu
North Kazakhstan	Ruzaevka	Akmola	Zhaksy	Kostanay	Presnogor'kovka
North Kazakhstan	Vozvyshenka			Kostanay	Tobol
				Kostanay	Zheleznodorozhnyy

Probes used in the Regression Equations dureing the month of June, July & August

Oblast	Probe Stations	Oblast	Probe Stations	Oblast	Probe Stations
North Kazakhstan	Blagoveschenka	Akmola	Arshaly	Kostanay	Arshalinskiy
North Kazakhstan	Sergeevka	Akmola	Atbasar	Kostanay	Diyevskaya
North Kazakhstan	Vozvyshenka	Akmola	Balkashino	Kostanay	Dzhetygora
North Kazakhstan	Blagoveschenka	Akmola	Novoishimsky	Kostanay	Fedorovka
North Kazakhstan	Chkalovo	Akmola	Voznesenka	Kostanay	Karasu
North Kazakhstan	Vozvyshenka	Akmola	Zhaltyr	Kostanay	Kostanay
North Kazakhstan	Chkalovo			Kostanay	Mikhaylovka
North Kazakhstan	Taiynsha			Kostanay	Tobol
North Kazakhstan	Vozvyshenka			Kostanay	Arshalinskiy

Figure 2. The probes stations used in the study, organized by oblast and period.

The PD (dependent variable) was regressed on the WI (independent variable) for each station and monthly time step. Then, independently for each oblast, a combination of the probe stations was combined to generate to regression equation for each monthly set. Months shared similar model parameters were combined into composite models. This led to making composite models for June, July, and August, while May retained its unique model equations.

The goal was to identify an equation that promoted high precision and retain a high degree of freedom. Data was merged from all the three oblasts, which allowed identification of the true relationship between the two data sets and would provide greater confidence in predicting the soil moisture from the WI. It was apparent from the slope and intercept as well as regression equations that they came from the same distribution (population). Data could therefore be effectively merged into a more comprehensive model that could be used predict the probe observation from the satellite derived measurements.

4. DATA

There are numerous approaches to deriving soil moisture from *in situ* observations. Arya and Paris (1981) describe how to use models that combine physical and empirical measurements to detect the bulk distribution of moisture in the soil. Cambardella et al. (1994) uses empirical techniques to detect the spatial distribution in soil moisture across areas in Iowa. Haubrock, et al. (2008) developed a technique that employs hyperspectral and field observations to detect the quantity of moisture in the soil. Kodikara J. et al. (2014) demonstrates how neutron probes can identify the amount of water in the soil. Data was used from the Kazakhstan probe station network. The probe data is gathered by probe station technicians who use a tool to extract a profile of soil (e.g., plug) from a location designated at one of the observation sites. Then the plug of soil is weighed to obtain the total weight of soil and moisture. Following this procedure, the plug of soil is slowly heated to evaporate all the moisture available to plants. The plug of soil is then weighed again to obtain the volumetric weight of water that was held in the soil plug.

These *in situ* measurements have great value as direct observations of the soil moisture profile. Measurements contain three profiles: top 20 cm, top 50 cm, and top 100 cm of the soil. Since satellite observations most closely align to probe data of the top 20 cm of the soil and data extracted from the top 20 cm of the soil makes up the most complete data set, it was used in the analysis.

Because *in situ* measurements are single point observations, they may not effectively represent the surrounding area to which they are generally applied. Moreover, since precipitation and soil water holding capacity has a lot of spatial variability in a small area, the single measurement may not represent the true distribution of soil moisture in the larger area. Considering the challenges of monitoring soil moisture directly, a simple yet robust procedure is used to predict upper level soil moisture from surface wetness values derived from satellite data.

The Basist Wetness Index (BWI) was utilized. The BWI is a surface wetness index that ranges from zero, which represents no water detected near the surface, to a percentage of the radiating surface that is liquid water. Thus, the range is zero to 100, where 100 represents the entire surface is liquid water (Basist et al 2001). This index is derived from a linear relationship between channel measurements (Equation 1), where a channel measurement is the value observed at a particular frequency and polarization (i.e. the SSM/I observes seven channels).

[Equation 1] $$BWI = \Delta\varepsilon \cdot T_s = \beta_0[T_b(v_2) - T_b(v_1)] + \beta_1[T_b(v_3) - T_b(v_2)]$$

where the change of emissivity (Basist et al. 2001), $\Delta\varepsilon$, is empirically determined from global SSM/I measurements, T_s is surface temperature over wet or dry land, T_b is the satellite brightness temperature at a particular frequency (GHz), υ_n (n=1, 2, 3) is a frequency observed by the SSM/I instrument, β_0 and β_1 are estimated coefficients that correlate the relationship of the various channel measurements to observed surface temperature at the time of the satellite overpass. Specifically, the greater the wetness, the larger the differences between observed surface temperatures and observed channel measurements (Basist et al. 1998).

A model based on the PD was modeled to take advantage of the direct measurements. The goal is to use the PD as reliable calibration points, and the WI as a means to better understand the actual spatial distribution of upper level soil moisture across northern Kazakhstan. Another advantage of the satellite observations is that it allows for near real time monitoring. Therefore, if the relationship has value, the satellite will be able to provide current information on the spatial and temporal resolution of soil moisture throughout the wheat growing area of northern Kazakhstan.

5. RESULTS

The analyses below provide the relationships among the various regression equations. The summer months of June, July, and August are combined and related between the PD and WI in each oblast. This method shows the monthly relationships were fairly similar across the summer months and the equations have comprehensive applications. There were three models corresponding with May; one for each oblast. A description of the regression equations is presented in the following paragraphs.

First are results from the North Kazakhstan oblast during the summer period (Fig. 3). As demonstrated in figure 2, the explanatory power of the model is 67%, which means the model explains over two-thirds of the variance in the PD. It is interesting to note a slope of 5.5, which indicates that for every value change of 1 in the WI, there is a subsequent 5.5 grams increase in the probe measurements. Another important consideration is the intercept. Based on the equation, there are 7 grams of water in the profile at the time the Satellite would register there is "no observed water in the profile". Of course some of that water would be unavailable to the plants, since it would be held under considerable ionic bonds with soil particles.

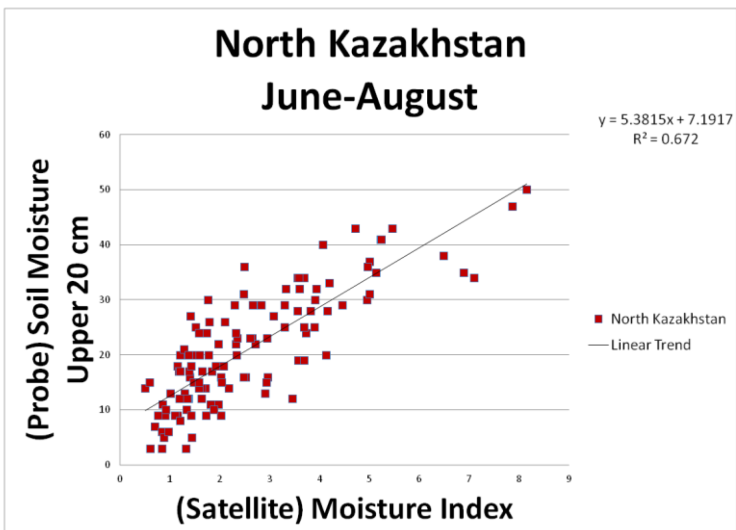

Fig. 3 The regression equation, where the Wetness Index predicts the Probe Data.

Based on the data provided from Akmola (Fig. 4) during the summer months, percent 52% of the variance in the PD was explained by the WI, which is slightly below the explanatory power of the Northern Kazakhstan model. Another point to note is a slope of 4.7 in the Akmola model, which is quite similar to the slope of the Northern Kazakhstan regression model. Moreover, the intercept of 6.1 is also quite similar between oblasts. These results lend support to the stability of the models and the similarity in their parameters. These results are positive and demonstrate model stability.

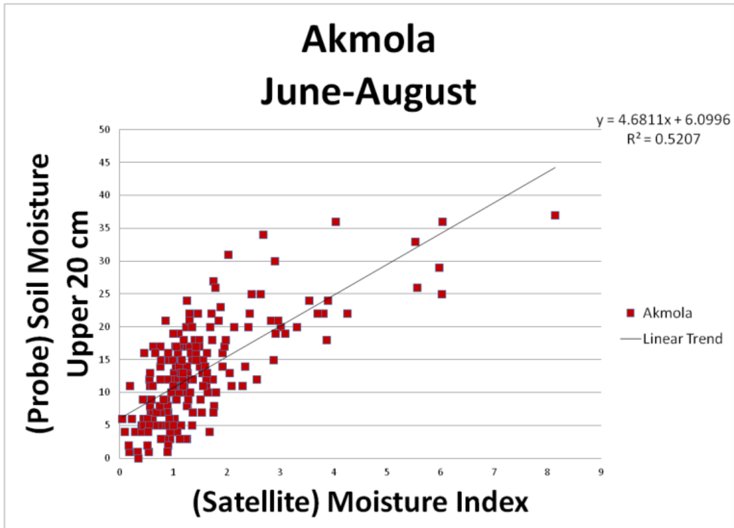

Fig. 4 The regression equation, where the Wetness Index predicts the Probe Data.

The Kostanay model during the summer months (Fig.5) has an explanatory power of 50%. Although it is slightly lower than the previously two oblasts, its skill is not appreciably less. The slope of the regression model is 5.7, which is slightly higher than Northern Kazakhstan and appreciably higher than Akmola. In terms of a *y*-intercept, one would expect it to be lower than the other two oblasts since its slope is higher. Indeed, that is the case; it has a value of 4.6. Reviewing these models as a group, it is apparent that they have many similarities and generally they represent the same relationship between the PD and WI.

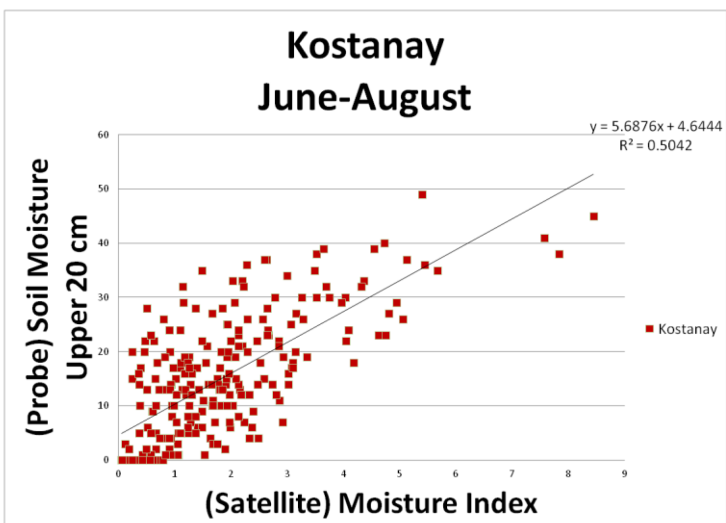

Fig. 5 The regression equation, where the Wetness Index predicts the Probe Data.

To generalize the applications of the three models, one could state that when the WI goes to zero, there is less than 5 to 7 grams of liquid water in the top 20 cm of the soil profile. The relationship between the PD and WI generally has a 5:1 ratio. Moreover all three models were able to explain over 50% of the variance in the PD. The stability of the model and their accuracy shows that the WI could be useful for identifying the availability and variability of upper level soil moisture during the summer months (in these three oblasts of northern Kazakhstan).

The next discussion will focus on the relationship between the PD and the WI during the month of May. Again, the relationships were derived from the period of record (10 years). A reference to figure 2 will identify which probe stations were retained in the formulation of these regression equations. It is apparent that fewer stations were utilized, since there is only one month of data and the early years were largely devoid of data. Despite this limitation, meaningful findings from the analyses were derived.

The Akmola model (Fig. 6) has a lower explanatory power than the models generated from data over the summer months. None-the-less, the Akmola regression equations had an explanatory power of 47%. The slope was 2.7, which is appreciably lower than the model from the summer months. The cause for this difference is not apparent, but it could be related to a higher percentage of the radiating surface that is bare soil, while the summer surface is vegetated. In other words, vegetation has an appreciable influence on the radiating surface.

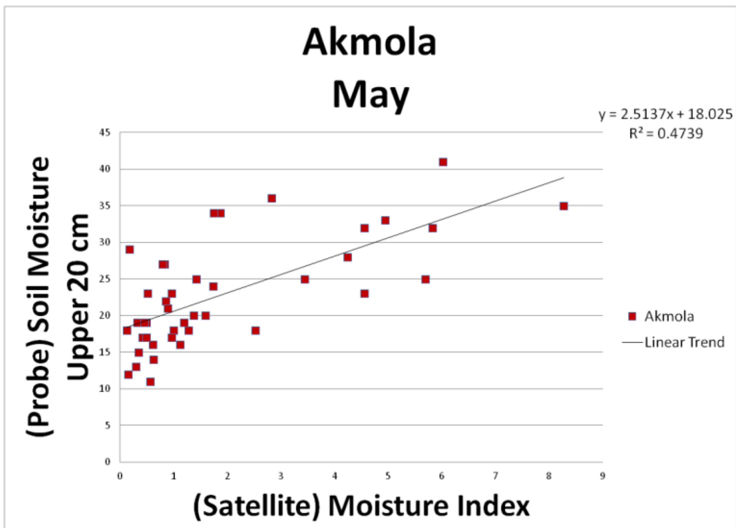

Fig. 6 The regression equation, where the Wetness Index predicts the Probe Data.

The regression equation for May in North Kazakhstan (Fig. 7) had an explanatory power of 22%, which is much lower than the previous models discussed above. Consequently, its predictive skill was only significant at the 0.05 confidence level. The slope of the equation was 1.6, which is considerable less than the prior model slopes. As a consequence of the low slope values, there is a much higher intercept at 28 grams of moisture, which corresponds to a zero value in the WI. This model was dismissed as it does not promote much confidence.

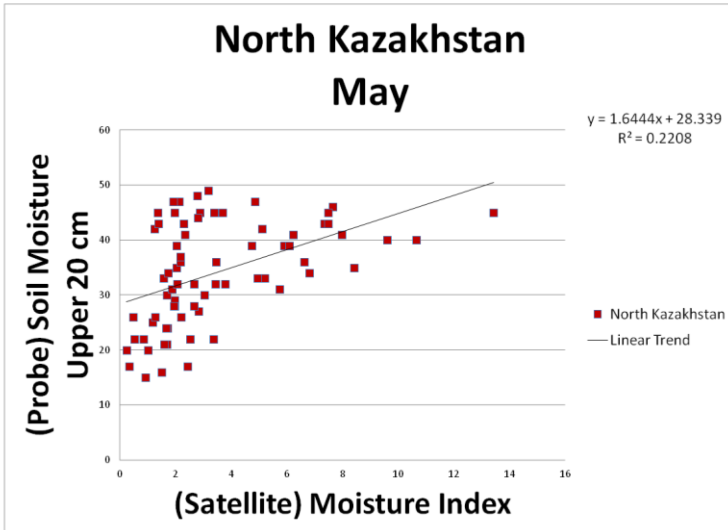

Fig. 7 The regression equation, where the Wetness Index predicts the Probe Data.

The regression model for Kostanay in May (Fig. 8) had a prediction skill of 62%, which is one the best explanatory powers of the six models. Moreover, the slope of this regression equation is 2.8, which is very close the slope of the Akmola model for the month of May. This similarity is also true for the two similar intercepts. These results indicate that one comprehensive model (based on these two oblasts, during the month of May) could provide a stable and realistic relationship between the PD and WI.

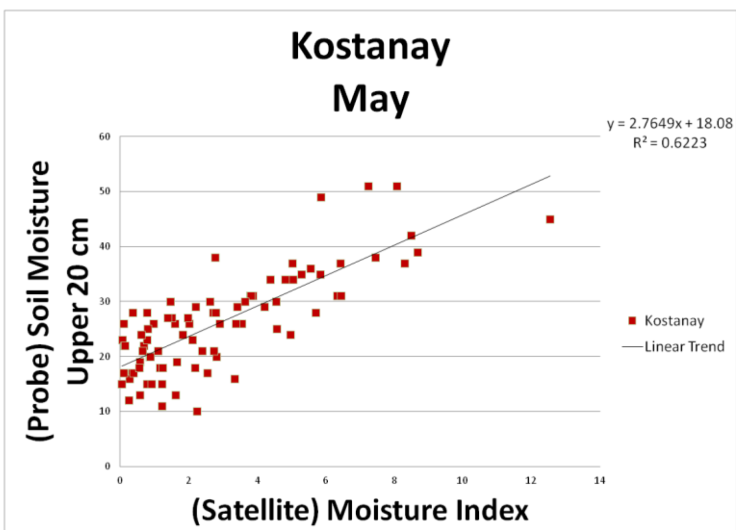

Fig. 8 The regression equation, where the Wetness Index predicts the Probe Data.

The May data for Akmola and Kostanay were combined into one comprehensive model. All three oblasts for the summer months were also combined into a different comprehensive model. The justification of the integration of the data sets into their respective models is based on similar: 1. Slopes, 2. Intercepts, 3. and Population distribution. Since the goal was to identify a comprehensive model, an integrated regression equation that may define a clear relationship between the PD and WI was chosen.

The summer time comprehensive model (Fig. 9) had a high explanatory power of 57%. Usually this decreases as sample areas increase. An advantage of the larger sample is that it does not over specify a regression equation from a limited number of observations. This makes the models better at predicting. The Slope was 5.5, which is a good compromise from the three equations integrated into one. The Intercept was also encouraging, since it fit in a tight range from the three data sets. The intercept shows a very low value of probe moisture, as the WI identified the surface as completely dry. This final model is a good approximation of the true relationship between the PD and WI, and that it has application in both near real time, and historical analysis. Consequently, the application of the model to access soil moisture quantities and distribution is promising and can be used to monitor recent moisture in the soil across northern Kazakhstan.

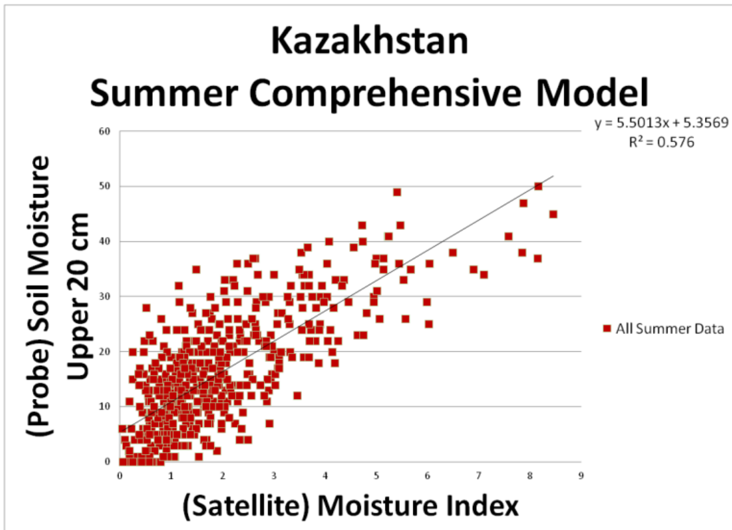

Fig. 9 The regression equation, where the Wetness Index predicts the Probe Data.

In terms of developing a comprehensive model for May, the data from Akmola and Kostanay was chosen due to appreciably higher R-square than North Kazakhstan. The comprehensive May regression equation had the explanatory power to 59%, which is favorable. Since both the Akmola and Kostanay equations had very similar slope and *y*-intercept, their integration into one model leans confidence that they converge towards the true relationship between PD and WI during the month of May.

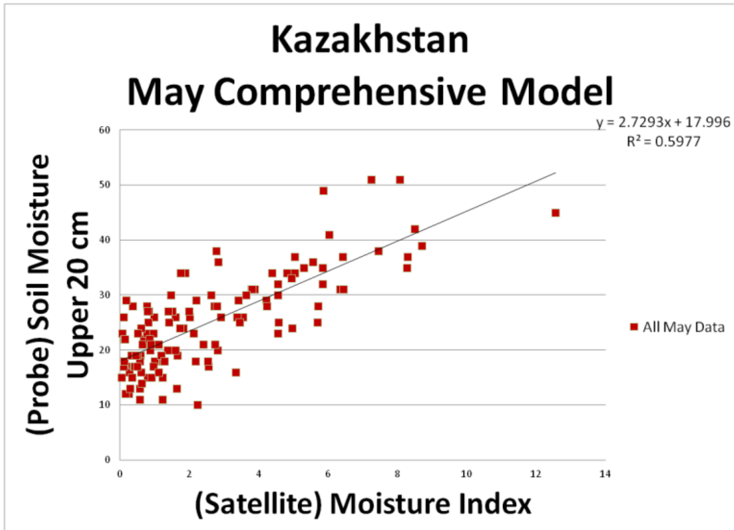

Fig. 10 The regression equation, where the Wetness Index predicts the Probe Data.

6. CONCLUDING DISCUSSION

This study used soil moisture from probe data (PD) provided by Kazhydromet and attempted to calibrate these measurements against satellite observations of surface wetness (WI). The satellite data was derived from the Special Sensor Microwave Imager (SSMI), which detects liquid water near the surface. Similar to Jury, W., and H. Vaux (2005), an attempt was made to introduce science and technology to improve understanding, as well as monitor environmental issues that directly impact social welfare. The WI is the percentage of the radiating surface that is liquid water, where the PD is the weight of the water in the column that evaporated from the soil sample. The soil samples are generally taken every 10 days: 8^{th}, 18^{th}, and 28^{th} of the month. The WI is generally observed twice a day, and averaged into weekly data sets.

Regression models were developed between the PD and WI for the months of June, July, and August (summer months) during the years 2003 and 2012. Models for various probes inside the oblasts were combined together for each month. Then, the months were combined into one comprehensive model. This model shows a more comprehensive relationship between the PD and WI. In the summer there is generally a 5.5 to 1 ratio for the PD/WI and explanatory power of the model is 58%. This result shows that the comprehensive model has value in explaining the relationship between soil moisture derived from probe data to the wetness values derived from the satellite observations. Consequently, this regression equation and the WI can be used to perform near real assessment of soil moisture at both moderate resolutions and weekly time intervals during the important periods of the growing season for winter wheat.

The relationships for May had a similar explanatory skill (60%) to the summer months. However, in terms of slope they were significantly different (the slope was 2.7 and the intercept 18). As stated in above, generally when the slope is less the intercept is higher, which has consistently been true in these analyses. The findings demonstrate a clear relationship between the PD and WI during the month of May, which is significantly different than the summer months. The reason for these different relationships is not fully understood. However, it may be due to considerable portions of satellite derived wetness signal originating from vegetative cover during the summer months, whereas the vast majority of the signal comes directly from the soil during the month of May.

Further research is needed based on the positive findings discovered by this study. The natural next step is to implement the regression equations into software that would automatically generate the distribution of the soil moisture at the satellite resolution. Since the resolution is 33 by 33 km, the satellite would provide a much larger network of simulated probe data, i.e. of soil moisture in the upper 20 cm for the soil. Another advantage is that these data would be available in near real time, each week. There is a considerable amount of work to perform to generate that capability. The required effort would be beneficial to the Kazhydromet and stakeholders monitoring water resources and food security in central Asia.

LITERATURE CITED

Alsdorf, D, et al. 2007. "Measuring Surface Water from Space," *American Geophysical Union*, 45(2): 1-24.

Arya, Lalit M., and Jack F. Paris. 1981. "A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data." *Soil Science Society of America Journal* 45.6: 1023-1030

Basist, A., Grody, N. C., Peterson, T. C., and Williams, C. N. 1998. "Using the Special Sensor Microwave / Imager to Monitor Land Surface Temperatures, Wetness, and Snow Cover," *Journal of Applied Meteorology*, 37: 888-911.

Basist, A., C. Williams Jr., N. Grody, T.F. Ross, S. Shen, A. T.C. Chang, R. Ferraro, and M.J. Menne. 2001. "Using the Special Sensor Microwave imager to Monitor Surface Wetness," *Journal of Hydrometeorology*, 2: 297-308.

Cambardella, C. A et al. 1994: Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, *58*(5): 1501-1511.

Griffiths, G. H., & Wooding, M. G. (1996). Temporal monitoring of soil moisture using ERS-1 SAR data. *Hydrological Processes*, *10*(9): 1127-1138.

Haubrock, et al. (2008). Surface soil moisture quantification and validation based on hyperspectral data and field measurements. *Journal of Applied Remote Sensing*, *2*(1): 023552-023552.

Jury, W., and H. Vaux. 2005. "The role of science in solving the world's emerging water problems," *PNAS*, 102(44): 15715-15720.

Kodikara J. et al. 2014. Soil moisture monitoring at the field scale using neutron probe *Canadian Geotechnical Journal*, 51(3): 332-345.

Robinson D.A. et al. 2008: Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review Soil Science Society of America, v. 7 no. p. 358-389.

Vörösmarty, Charles, Pamela Green, Joseph Salisbury, and Richard Lammers. 2000. "Global Water Resources: Vulnerability from Climate Change and Population Growth," *Science*, 289: 284-288.

Wagner, et al. (2007). Soil moisture from operational meteorological satellites. *Hydrogeology Journal*, *15*(1), 121-131.

Weinthal, Erika. 2002. *State Making and Environmental Cooperation: Linking Domestic and International Politics in Central Asia.* Cambridge, MA: The MIT Press.

U.S. Agency for International Development

1300 Pennsylvania Avenue, NW

Washington, DC 20523

Tel: (202) 712-0000

Fax: (202) 216-3524

www.usaid.gov